Estimation of Probability Distributions for Individual Parameters Using Aggregate Population Observations

H.T. Banks

B.G. Fitzpatrick

Laura K. Potter

Yue Zhang

January 26, 1998

Dedicated to Wendell H. Fleming on the occasion of his 70th birthday.

Abstract

In this paper we discuss a general methodology for estimating the distribution of individual growth rates in a size-structured population using aggregate population data. The method, for which rigorous theoretical formulations have been developed, is presented in the context of an inverse problem methodology and its use is illustrated with application to mosquitofish, *Gambusia affinis*, population in rice fields.

1 Introduction

In this paper, we present results using inverse problem techniques for estimation of growth distribution in size-structured population models using aggregate population data. The models employed here are based on ideas initially discussed in [BBKW], which entail models wherein growth rates may vary across individuals of the population as well as with size and time.

These models are in contrast to the usual stochastic partial differential equation models as described for example in [FL1],[FL2],[FS] and in the next section. Although they are not stochastic in the usual sense, they are probabilistic in that one models growth, mortality, etc. via probability distributions across the population.

Simulation studies were presented in [BBKW] to demonstrate that such ideas could lead to population densities that exhibit dispersion and bimodality. Rigorous theoretical developments of the associated inverse problem technique and initial illustrations with computational examples were given in [BF], [F1] and [F2]. A survey of results and other references can be found in [B]. The general

1

1. REPORT DATE **26 JAN 1998**	2. REPORT TYPE	3. DATES COVERED **00-00-1998 to 00-00-1998**
4. TITLE AND SUBTITLE **Estimation of Probability Distributions for Individual Parameters Using Aggregate Population Observations**		5a. CONTRACT NUMBER
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **North Carolina State University,Center for Research in Scientific Computation,Raleigh,NC,27695-8205**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**		
13. SUPPLEMENTARY NOTES		
14. ABSTRACT **see report**		
15. SUBJECT TERMS		

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES **19**	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**			

philosophy underlying our approach of using aggregate population data to estimate individual rates is given (along with an application to susceptibility and vaccination efficiency in populations) in [BFZ].

As detailed in earlier references, our efforts on such problems were initiated in collaboration with marine biologists (specifically Lou Botsford and his colleagues) in studies related to the introduction of mosquitofish into rice fields, in attempts to control mosquito populations without chemicals. To establish viable control strategies, it is very important to have quantitative models which predict accurately the evolution of the populations. In the paper [BVWLKRC], the mosquitofish population was modeled using the Sinko-Streifer (also called the McKendrick-Von Foerster) model for size-structured population density evolution. As we shall discuss below, this leads to a number of conceptual difficulties.

2 Modeling Philosophy

There is a huge literature (we won't even attempt to cite these here) on the modeling of populations in the quantitatively oriented biological research literature. The efforts range from modeling population growth, death, mortality in insects, marine life, plants and animals to gene frequency mutation and drift, as well as treatments of susceptibility to disease in vaccinated populations of humans. In such a diverse scientific literature it is often difficult to discern an underlying commonality. However, there are some usual features in much of the literature. The modeling often relies on data about certain observables, while it is knowledge about other nonobservable parameters that is of interest to investigators. For example, one usually has total population (aggregate densities) counts in models where growth, mortality and migration rates of a typical individual are of most interest. Or one may have aggregate data on percent of a population vaccinated, and numbers of those who fall ill and recover (or not) in studies of disease prophylaxis. In the case of genetic studies one has data on phenotypes, where it is gene frequencies and their changes that one wishes to understand.

A second feature of most modeling attempts is the presence of uncertainty. This may arise in the model itself, in parameters in the model, in distribution over the population of unobservable traits or characteristics. A conceptually important question is how to treat these uncertainties when attempting to extract the maximum information about the population from the data. The standard approach to treat probabilistic or stochastic aspects of population is through the use of stochastic differential equations. A brief review of examples early on in this approach was given by Fleming in [FL1], where he discussed distributed parameter or partial differential equation models for geographically-structured (including size or age-structured) populations. He also discussed population genetics models (equations for the mean and the covariance of gene frequency in spatially migrating populations — see [FL2],[FS]). Our focus here will be on size-structured population models and our use of rate distribution models as an alternative to stochastic differential equations. To put this in context of the more standard approach, we introduce and discuss briefly the Fokker-Planck

size-structured model.

The Fokker-Planck equation is the basis of a stochastic size-structured model [B1],[BTW] which is based on the assumption that movement from one size class to another can be described by a Markov process. The "physiological age" interpretation of the Fokker-Planck equation was first suggested by Weiss [W]; a careful derivation based on the paradigm of Brownian motion of particles which is applicable to growth processes is given in [O]. The Fokker-Planck equation, under the assumption of a Markov transition for the growth process, is

$$\frac{\partial u}{\partial t}(t, x) + \frac{\partial}{\partial x}(\mathcal{M}_1(t, x)u(t, x)) = \frac{1}{2}\frac{\partial^2}{\partial x^2}(\mathcal{M}_2(t, x)u(t, x)), \tag{1}$$

where $u(t, x)$ is the population density at time t and size x and the moments are given by

$$\mathcal{M}_j(t, x) \equiv \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{-\infty}^{\infty} (y - x)^j p(t, x; t + \Delta t, y)dy. \tag{2}$$

The function $p(t, x; t + \Delta t, y)$ is the probability density for the transition from size x at time t to size y at time $t + \Delta t$; i.e., $p(t, x; t + \Delta t, y)\Delta x$ is the probability that members in the size interval $[x, x + \Delta x]$ at time t will move to size y at time $t + \Delta t$. The moments $\mathcal{M}_1$, $\mathcal{M}_2$ have probabilistic interpretations: $\mathcal{M}_1$ is the mean (or first moment) of the rate of increase in size

$$\mathcal{M}_1(t, x) = \lim_{\Delta t \to 0} \mathcal{E}\left\{\frac{x(t + \Delta t) - x(t)}{\Delta t}\right\},$$

where $\mathcal{E}[X]$ denotes the expected value of a random variable X, while $\mathcal{M}_2$ is the second moment of the rate of increase in size

$$\mathcal{M}_2(t, x) = \lim_{\Delta t \to 0} \mathcal{E}\left\{\frac{(x(t + \Delta t) - x(t))^2}{\Delta t}\right\}.$$

Appropriate boundary conditions must be formulated for (1). Since $\mathcal{M}_1 u - \frac{\partial}{\partial x}\left(\frac{\mathcal{M}_2}{2}u\right)$ represents the population flux, we have

$$\left[\mathcal{M}_1(t, x)u(t, x) - \frac{\partial}{\partial x}(\mathcal{M}_2(t, x)u(t, x))\right]^{x=x_0}$$

$$= \int_{x_o}^{x_m} K(t, x)u(t, x)dx \tag{3}$$

$$\left[\mathcal{M}_1(t, x)u(t, x) - \frac{\partial}{\partial x}(\mathcal{M}_2(t, x)u(t, x))\right]^{x=x_m} = 0 \tag{4}$$

while the initial conditions are given by

$$u(t_0, x) = \Phi(x). \tag{5}$$

Here x_m is maximum size, x_0 is the minimum size, and K is the fecundity rate as explained below in our discussion of the Sinko-Streifer model.

The system (1), (3)-(5) comprises an initial-boundary value problem for a transport-dominated diffusion process (the $\mathcal{M}_1$ term is typically dominant over the $\mathcal{M}_2$ term) that offers significant computational challenges. First, the moments $\mathcal{M}_1, \mathcal{M}_2$ almost always (for the applications to populations) depend on both t and x and must be estimated; the expressions (2) are not helpful since, in general, p is unknown. Moreover, even simulation with (1), (3)-(5) is nontrivial. Traditional finite difference and finite element methods produce erroneous oscillatory solutions. Thus the spline-based, fixed node methods first proposed in [B1] are of very limited interest. However, high promise can be found [BTW] in a moving node finite element technique first suggested in [H]. Even so, estimation or inverse problems built on Fokker-Planck models are extremely different computationally and have enjoyed limited use in the literature.

Our efforts to treat both unobservable individual parameters as well as a certain amount of stochasticity is built upon the Sinko-Streifer model and thus we briefly outline its features.

The classical Sinko-Streifer model, henceforth referred to simply as (SS), is given by

$$\frac{\partial}{\partial t}v(t,x) + \frac{\partial}{\partial x}(g(t,x)v(t,x)) = -\mu(t,x)v(t,x) \qquad t > t_0 \ ,$$

$$v(0,x) = \Phi(x) \qquad\qquad x_0 < x < x_m \ , \tag{6}$$

$$g(t,x_0)v(t,x_0) = \int_{x_0}^{x_m} K(t,\xi)v(t,\xi)d\xi$$

$$g(t,x_m) = 0$$

and simulates the time evolution of a population with respect to the size x of the individuals. In our use of this model, the parameter $x \in [x_0, x_m]$ denotes the length of the fish, and the function v represents the size density function so that

$$N(t) = \int_a^b v(t,x)dx \tag{7}$$

is the number of fish in the population at time t whose size is between a and b. The function g represents the growth rate of the individuals,

$$\frac{dx(t)}{dt} = g(t,x) \ ,$$

so that in this simple model all individuals have the same growth rate. The last condition in (6), $g(t,x_m) = 0$, indicates that growth ceases when individuals reach maximum size x_m. The initial condition Φ describes the initial size density for the population. The function K is a fecundity kernel, and is used to express the recruitment rate $R(t,v) = \int_{x_0}^{x_m} K(t,\xi)v(t,\xi)d\xi$. The mortality rate is given by μ.

Difficulties arise, however, in applying the model (SS) to observed data since the data often exhibit features that the model cannot simulate or predict. One such phenomenon is a dispersion in size as time progresses. Another is that the population begins with a unimodal density and in time develops into one with

4

a bimodal density. Both of these features are present in the mosquitofish data (see [BBKW], [BVWLKRC], and the data plot in Figure 1).

This dispersion in size and the transition from a unimodal to a bimodal density is qualitatively inconsistent with solutions of the Sinko-Streifer model under biologically feasible assumptions on the model. We are able to capture both phenomena with growth rate distribution (GRD) models.

2.1   GRD Models with Identical Initial Size Densities

Here we use a modification of the model (SS) (in actuality, a continuum of (SS) across families of growth rates) which exhibits features present in the data. We use as the aggregate population density (APD) the function

$$u(t, x) = \int_G v(t, x; g) dP(g), \tag{8}$$

where G is a collection of growth rates and P is a probability measure on G. This approach was first suggested in [BBKW]. As detailed in [BF], the resulting model is rich enough to exhibit the phenomena of interest, namely, dispersion and development of two modes from one. As also explained in [BF], measure theoretic results allow us to approximate the continuum measure in (8) by a discrete measure corresponding to a finite dimensional set of growth rates.

Briefly, we assume that we have a family $G = \{g_{jk}\}$, $j = 1, \cdots, M_1$, $k = 1, \cdots, M_2$ of individual growth rates; individuals are grouped into the same subpopulation if they have the same growth rate. The subpopulations grow according to the (SS) with $g = g_{jk}$ in (6). Here we take $g_{jk}(x) = r_j(\gamma_k - x)$, $\mu = 0$ and $K = 0$. Note, however, that a generalization to allow distribution of mortality and fecundity over subpopulations is readily achieved.

The size density for the subpopulation jk is given by $v(t, x; g_{jk})$, and the aggregate population density is given by

$$u(t, x) = \sum_{j,k} v(t, x; g_{jk}) p_{jk}, \tag{9}$$

where p_{jk} is the probability of an individual being in the jk subpopulation. Note that this formulation embodies the assumption that each subpopulation has the same initial size density, i.e., $v(0, x; g_{jk}) = \Phi(x)$ for all j, k.

We used the above outlined formulation in a least squares inverse problem involving fitting several sets of field data $\{\hat{u}(t, x)\}$. The probabilities $\{p_{jk}\}$ for $1 \leq j \leq M_1$ and $1 \leq k \leq M_2$ in (9) were estimated by solving the following inverse problem:

$$\min_{\mathbf{p} \in P^M(G)} J(\mathbf{p}) \quad = \quad \sum_{n=1}^{N} \int_{x_0}^{x_1} [\sum_{j,k} v(t_n, x; g_{jk}) p_{jk} - \hat{u}(t_n, x)]^2 dx \tag{10}$$

$$= \quad \sum_{n=1}^{N} \int_{x_0}^{x_1} ([\sum_{j,k} v(t_n, x; g_{jk}) p_{jk}]^2$$

$$-2\hat{u}(t_n, x) \sum_{j,k} v(t_n, x; g_{jk}) p_{jk} + [\hat{u}(t_n, x)]^2 dx.$$

Here $P^M(G)$ is a finite dimensional approximation to the probability measure space $P(G)$ defined by

$$P^M(G) = \{P \in P(G) | P = \sum_{j,k} p_{jk} \delta_{g_{jk}}, \quad \sum_{j,k} p_{jk} = 1\}$$

where $M = M_1 \times M_2$ and $\delta_{g_{jk}}$ is the Dirac measure with an atom at g_{jk}. We denote by $\mathbf{p}$ the array that contains p_{jk}, $1 \le j \le M_1$, $1 \le k \le M_2$, and set

$$a_{jk\ell m} = \sum_{n=1}^{N} \int_{x_0}^{x_1} v(t_n, x; g_{jk}) v(t_n, x; g_{\ell m}) dx$$

$$b_{jk} = -\sum_{n=1}^{N} \int_{x_0}^{x_1} \hat{u}(t_n, x) v(t_n, x; g_{jk}) dx$$

$$c = \sum_{n=1}^{N} \int_{x_0}^{x_1} [\hat{u}(t_n, x)]^2 dx .$$

Moreover, we define $\mathbf{A}$ and $\mathbf{b}$ to be the arrays that contain $a_{jk\ell m}$ and b_{jk} respectively, for $1 \le j, \ell \le M_1$ and $1 \le k, m \le M_2$.

The inverse problem (10) now becomes a quadratic programming problem in which one minimizes $\mathbf{p}^T \mathbf{A} \mathbf{p} + 2 \mathbf{p}^T \mathbf{b} + c$ over $P^M(G)$. To include the equality constraint, $\sum p_{jk} = 1$, we introduce the Lagrange multiplier λ and solve the unconstrained problem $\min F(\mathbf{p}, \lambda)$, where

$$F(\mathbf{p}, \lambda) \equiv \mathbf{p}^T \mathbf{A} \mathbf{p} + 2 \mathbf{p}^T \mathbf{b} + c + \lambda [\sum_{j,k} p_{jk} - 1] . \tag{11}$$

We used the "method of characteristics" based techniques discussed in [IKP] to solve (6) for the densities $v(t, x; g_{jk})$, which were calculated in parallel using PVM software on six IBM RISC 6000 workstations. These densities were then used with the IMSL subroutine $DQPROG$ to solve for the optimal solution $(\mathbf{p}^*, \lambda^*)$. Finally, the aggregate population density $u(t, x)$ was determined by substituting p_{jk}^* into (9).

In the above parameter estimation problem, for each subgroup jk, a proportioned initial size density $\Phi(x)p_{jk}$ was assumed. The Day 195 data (see Figure 1) was chosen so that

$$\sum_{j,k} v(t_0, x; g_{jk})p_{jk} = \sum_{j,k} \Phi(x)p_{jk}.$$

Note that since the first data set is used here as the initial size density, it cannot be used in solving the inverse problem. This leaves one less data set available for use in estimating the optimal parameter set.

Example 1: We present in the following an example which is typical of the field-data fitting results. The fish were stocked on June 28, 1982 into four rice paddies with parallel water flow. Each paddy had an outflow trap to measure emigration. Measurements were taken weekly (two paddies one day and two the next) with the use of fifteen traps per paddy. The total number of fish caught was greater than or equal to the number actually measured.

The size distribution frequency for size class i is defined as $f_i = n_{m,i}/N_m$, where $n_{m,i}$ is the number of fish measured in size class i, and N_m is the total number of fish measured.

The total population was divided into five hundred and twelve subpopulations with

$$r_j = 0.2 + 1/31 * 4.8 * (j-1) \qquad j = 1, 2, ..., 32$$
$$\gamma_1 = 16/38, \ \gamma_2 = 22/38, \ \gamma_3 = 24/38,$$
$$\gamma_k = 16/38 + 1/15 * 22/38 * (k-1) \qquad k = 4, ..., 16.$$

The range for r_j was chosen arbitrarily, while the γ_k were based on the field data. The choice for $\gamma_1 = 16/38$ was made by inspecting the data from Day 195. At $x = 16$mm, a significant decrease in the number of fish occurs, which we interpreted as the smallest possible maximum length for some subpopulation. We obtained $\gamma_1 = 16/38$ and $\gamma_{16} = 1.0$ after normalizing against the largest possible length, $x_1 = 38mm$, for the total population. The initial size density $\Phi(x)$ (for all subpopulations) was approximated by interpolating the Day 195 data.

The computed results manifest dispersion and bimodality as we expected (see Figure 1), with a residual of $J_1 = 3.7626 \times 10^{-4}$. Figures 2 and 3 illustrate the probability densities and distribution respectively as a function of r and γ.

Although the computed results do provide dispersion and bimodality, the mismatches between the computed solutions and the field data are still significant. Furthermore, the discrepancies did not diminish as fast as hoped when we increased the number of subpopulations, $M = M_1 \times M_2$. The idea of varying the initial size density among the subgroups became appealing.

2.2 GRD Models with Subpopulation Dependent Initial Size Densities

A further refinement (and one which is appropriate for the mosquitofish population presented here) entails a parameterization of the initial density
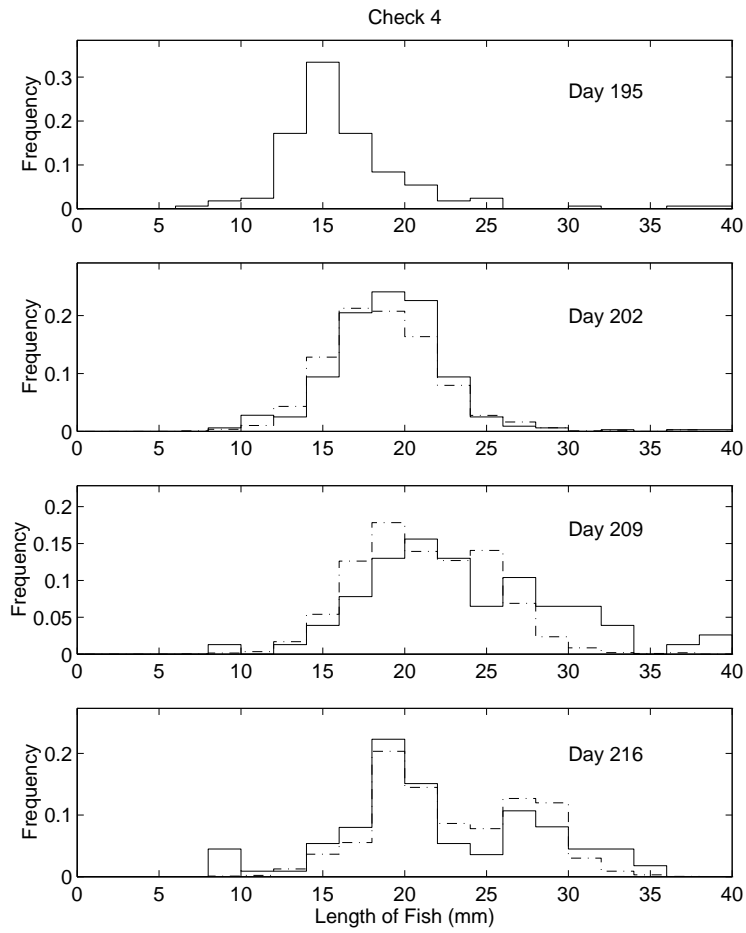
Figure 1: Computed results (- .) vs. field data (—) for Example 1.
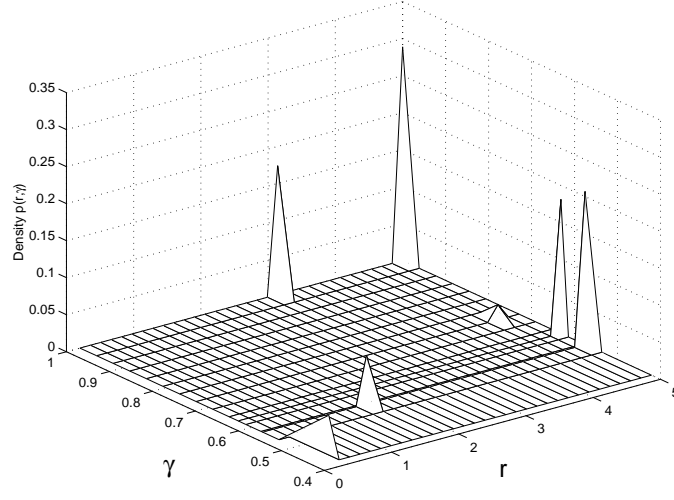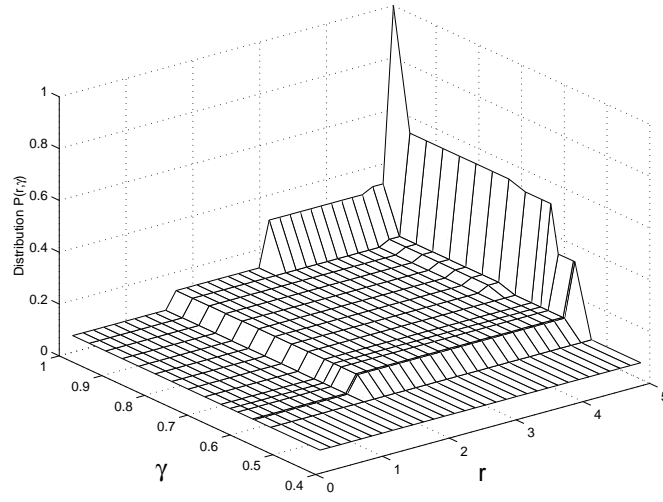
Figure 2: Probability densities p_{jk} for Example 1.

Figure 3: Probability distribution P_{jk} for Example 1.

$\Phi(x) = \sum_{j,k} \phi_{jk}(x)$ with $\phi_{jk}(x) = \sum_\ell \chi_\ell(x)\alpha_{\ell;j,k}$. Here $\chi_\ell(x)$ is the characteristic function of the interval $[x_\ell, x_{\ell+1})$ corresponding to a partition $0 = x_0 = x_1 < x_2 < \cdots < x_{M_3+1} = 1$, and we have normalized the size range $[x_0, x_m]$ to [0,1]. If one solves (6) with $g = g_{jk}$, $\Phi = \chi_\ell$, then $v(t, x; g_{jk}, \chi_\ell)$ represents the density for individuals with growth rate g_{jk} that initially have size structure density given by χ_ℓ.

The subpopulation jk density is given by

$$v(t, x; g_{jk}) = \sum_\ell v(t, x; g_{jk}, \chi_\ell)\alpha_{\ell;j,k}$$

so that the aggregate population density is then given by

$$u(t, x) = \sum_{j,k} v(t, x; g_{jk})p_{jk} = \sum_\ell \sum_{j,k} v(t, x; g_{jk}, \chi_\ell)p_{jk}\alpha_{\ell;j,k}. \qquad (12)$$

As before, we require the constraint $\sum_{j,k} p_{jk} = 1$. We denote α as the array containing $\alpha_{\ell;j,k}$ for $1 \le j \le M_1$, $1 \le k \le M_2$ and $1 \le \ell \le M_3$.

This formulation leads to a new inverse least squares problem with objective function

$$
\begin{aligned}
J(\mathbf{p}, \alpha) &= \sum_{n=0}^{N} \int_0^1 |u(t_n, x) - \hat{u}(t_n, x)|^2 \, dx \qquad (13) \\
&= \sum_{n=0}^{N} \int_0^1 \left| \sum_{j,k} \sum_{\ell=1}^{M_3} v(t_n, x; g_{jk}, \chi_\ell)p_{jk}\alpha_{\ell;j,k} - \hat{u}(t_n, x) \right|^2 dx
\end{aligned}
$$

subject to $\sum_{j,k} p_{jk} = 1$, and $v(t, x; g_{jk}, \chi_\ell)$ being the solution of (6) corresponding to $g = g_{jk}$ and $\Phi = \chi_\ell$. Minimizing (13) over both $\mathbf{p}$ and α requires solving a nonlinear programming problem.

An alternative and simpler approach to directly solving (13) is to reduce the nonlinear programming problem back to a quadratic programming problem. This can be accomplished by eliminating either $\mathbf{p}$ or α from the minimization process.

2.2.1 Minimizing Over α Only

One method for reducing (13) to a quadratic programming problem is to remove $\mathbf{p}$ from the parameter estimation problem and simply minimize over α. That is, if we can pre-determine a good choice for $\mathbf{p}$, we can fix the probabilities p_{jk} in (13) and reformulate it as a quadratic programming problem with objective function

$$F(\alpha) \equiv \alpha^{\mathrm{T}} \mathbf{A} \alpha + 2\alpha^{\mathrm{T}} \mathbf{b} + c. \qquad (14)$$

Here $\mathbf{A}$ is the $M_1 M_2 M_3 \times M_1 M_2 M_3$ array containing the elements

$$a_{jk\ell,qst} = \sum_{n=0}^{N} \int_0^1 [v(t_n, x; g_{jk}, \chi_\ell)p_{jk}][v(t_n, x; g_{qs}, \chi_t)p_{qs}]dx, \qquad (15)$$

b is the vector with elements

$$b_{jk\ell} = \sum_{n=0}^{N} \int_0^1 \hat{u}(t_n, x) v(t_n, x; g_{jk}, \chi_\ell) p_{jk} dx, \qquad (16)$$

and

$$c = \sum_{n=0}^{N} \int_0^1 [\hat{u}(t_n, x)]^2 dx. \qquad (17)$$

A good way to pre-determine the probabilities p_{jk} is to first solve the problem for (11) with identical initial size densities. We can then use the solutions $\mathbf{p}^*$ as the fixed probabilities in the arrays $\mathbf{A}$ and $\mathbf{b}$.

A major advantage of this approach is that the formulation allows for different initial size densities across subpopulations. Moreover, since the initial size densities are not estimated directly with the first data set, all of the data sets may be used in solving the inverse problem.

As before, we used "method of characteristics" based techniques, parallel computation, and the IMSL subroutine $DQPROG$ to solve for the optimal solution α^*. The resulting $\alpha^*_{\ell;j,k}$ and the pre-determined p^*_{jk} were then substituted into (12) to solve for $u(t, x)$ corresponding to $(\mathbf{p}^*, \alpha^*)$.

Example 2: In the following example, we use the 1982 field data as described in Example 1. First the probability measures p_{jk} were obtained by optimizing (11), where the total population was divided into thirty-two subpopulations according to the following:

$$r_j = 0.2 + 1/7 * 4.8 * (j-1) \qquad j = 1, 2, \ldots, 8$$

$$\gamma_1 = 16/38, \ \gamma_2 = 22/38, \ \gamma_3 = 24/38, \ \gamma_4 = 1. \qquad (18)$$

As in Example 1, the Day 195 data was used as the initial size distribution $\Phi(x)$ for all subgroups.

The computed probabilities p^*_{jk} were then used in solving the reformulated quadratic programming problem for (14) with subpopulation-dependent initial size densities. The entire population was further divided into a total of six hundred eight subpopulations, with r_j and γ_k as in (18) for $j = 1, 2, \ldots, 8$, $k = 1, 2, 3, 4$, and

$$\chi_\ell = [x_\ell, x_{\ell+1}) = [1/19 * (\ell - 1), 1/19 * \ell) \qquad \ell = 1, \ldots, 19.$$

The intervals χ_ℓ each correspond to a 2mm size class as seen in the field data.

The computed weights $\alpha^*_{\ell;j,k}$ were then substituted into (12) along with the precomputed p^*_{jk} to obtain the aggregate population density $u(t, x)$ (see Figure 4). This calculated density gave a better fit to the field data than the fit obtained in Example 1. Figures 5 and 6 illustrate the probability densities and distribution respectively as a function of r and γ.

The overall residual was $J_2 = 3.08780 \times 10^{-4}$. Since the parameter estimation carried out in Example 1 did not utilize the Day 195 data, we define a second
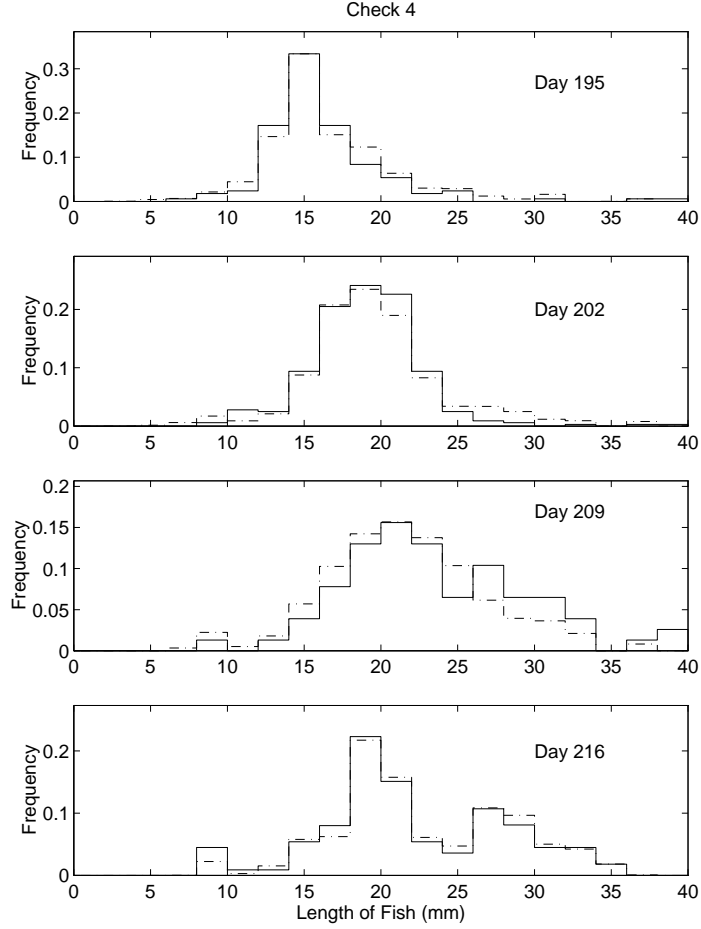
11

Figure 4: Computed results (- .) vs. field data (——) for Example 2.

residual $\tilde{J}_2$ which is taken only over Days 202, 209 and 216 of the data. This provides a means of comparing the accuracy of Example 1 versus Example 2.

In this manner we obtained $\tilde{J}_2 = 2.9305 \times 10^{-5}$, which is quite an improvement over $J_1 = 3.7626 \times 10^{-4}$.

2.2.2 Iterative Quadratic Programming

An alternative approach for simplifying the nonlinear programming problem for (13) is an iterative process involving two separate quadratic programming problems. The first of these quadratic programming problems is formulated by minimizing (13) over p_{jk} while $\alpha_{\ell;j,k}$ is held constant, and the second minimizes over $\alpha_{\ell;j,k}$ while p_{jk} is held constant. These quadratic programming problems
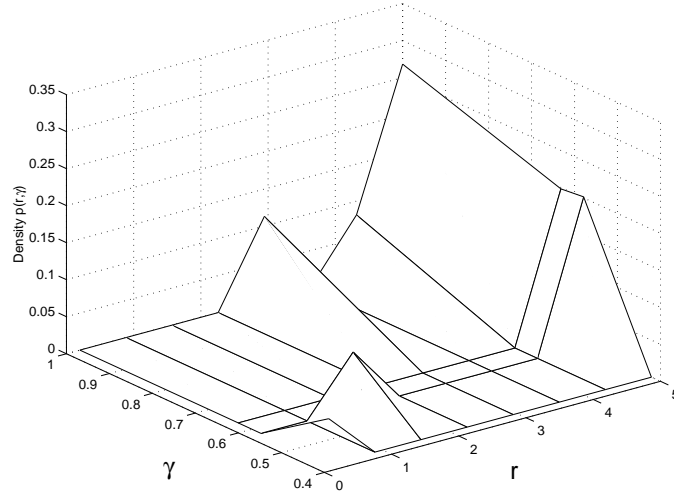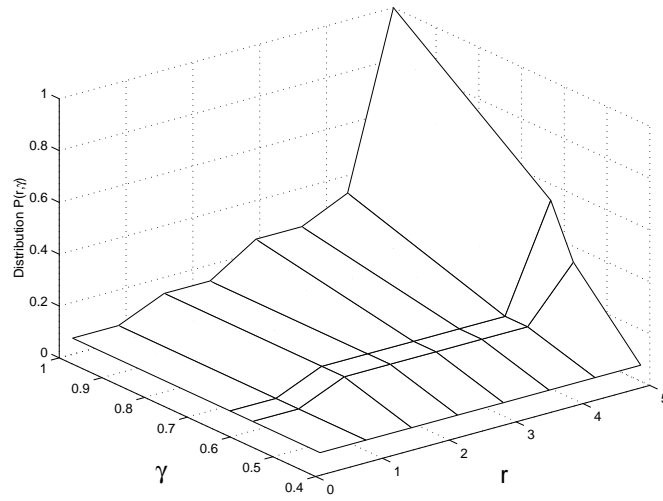
Figure 5: Probability densities p_{jk} for Example 2.

Figure 6: Probability distribution P_{jk} for Example 2.

can be solved in succession, creating an iterative process that progressively improves upon $\mathbf{p}$ and α.

Since the variables $\alpha_{\ell;j,k}$ are associated with the initial size densities ϕ_{jk}, we can approximate $\alpha_{\ell;j,k}$ using the initial time data. That is, we let

$$\chi_\ell = [x_\ell, x_{\ell+1}) \qquad \ell = 1, \ldots, M_3,$$

subdividing the size classes in the same manner as the data. For each ℓ, we then set

$$\alpha_{\ell;j,k} = \hat{u}(0, x_\ell)$$

for every j, k.

These values of $\alpha_{\ell;j,k}$ can then be inserted as fixed quantities in (13) to arrive at a quadratic programming problem with objective function

$$F(\mathbf{p}, \lambda) \equiv \mathbf{p}^T \mathbf{A} \mathbf{p} + 2\mathbf{p}^T \mathbf{b} + c + \lambda[\sum_{j,k} p_{jk} - 1]. \tag{19}$$

Here $\mathbf{A}$ is the $M_1 M_2 \times M_1 M_2$ array with elements

$$a_{jk\ell m} = \sum_{n=0}^{N} \int_0^1 [\sum_{i=1}^{M_3} v(t_n, x; g_{jk}, \chi_i)\alpha_{i;j,k}][\sum_{i=1}^{M_3} v(t_n, x; g_{\ell m}, \chi_i)\alpha_{i;\ell,m}]dx,$$

the vector $\mathbf{b}$ contains elements

$$b_{jk} = \sum_{n=0}^{N} \int_0^1 \hat{u}(t_n, x)[\sum_{i=1}^{M_3} v(t_n, x; g_{jk}, \chi_i)\alpha_{i;j,k}]dx,$$

and

$$c = \sum_{n=0}^{N} \int_0^1 [\hat{u}(t_n, x)]^2 dx.$$

Again we used the "method of characteristics" based techniques, parallel computation and the IMSL subroutine $DQPROG$ to obtain the optimal $(\mathbf{p}^*, \lambda^*)$.

These probability measures p_{jk}^* may now in turn be used to solve a quadratic programming problem over $\alpha_{\ell;j,k}$. In fact, by holding $\mathbf{p}$ constant in the original problem for (13), we have reduced the nonlinear programming problem to the quadratic programming problem involving (14)-(17). From this we obtain the optimal α^*.

This process can be repeated, leading to the following iterative method:

1. Partition the size interval into subintervals $\{x_\ell, x_{\ell+1}\}$, $\ell = 1, 2, \ldots, M_3$, so that the subintervals correspond to the subdivisions of size classes in the data. For each ℓ, let $\alpha_{\ell;j,k}^0 = \hat{u}(0, x_\ell)$.

2. (a) For $i = 0, 1, \ldots$, let $\alpha_{\ell;j,k} = \alpha_{\ell;j,k}^i$ and solve the quadratic programming problem for (19) to obtain the optimal p_{jk}^i.

14

(b) With $p_{jk} = p_{jk}^i$ for all j, k, solve the quadratic programming problem for (14) to obtain the optimal $\alpha_{\ell;j,k}^{i+1}$.

3. Stop iterating when $|\mathbf{p}^{i+1} - \mathbf{p}^i| < TOL_1$ and $|u^{i+1}(0, x) - u^i(0, x)| < TOL_2$, for some user-specified tolerances TOL_1 and TOL_2.

Our choice of the above stopping criteria was motivated by the structure of our model. The p_{jk} are probabilities associated with the subpopulations g_{jk}, and enter the model through (9). We would like $|\mathbf{p}^i - \mathbf{p}^{i+1}|$ to be within a specified tolerance. An additional stopping criterion should involve the variables $\alpha_{\ell;j,k}$, which are weights associated with the parameterization of the initial size density $\Phi(x) = \sum_{j,k} \sum_\ell \chi_\ell(x) \alpha_{\ell;j,k}$. The individual $\alpha_{\ell;j,k}$ themselves need not converge upon iteration, but we would like the initial size densities to converge. That is, we would like
$|u^i(0, x) - u^{i+1}(0, x)| < TOL_2$, for some specified tolerance TOL_2.

Example 3: In this example we continue to use the 1982 field data as detailed in Example 1. To initialize the iterative process, we approximated the weights $\alpha_{\ell;j,k}^0$ by the initial data from Day 195. That is, we let

$$\chi_\ell = [x_\ell, x_{\ell+1}) = [1/19 * (\ell - 1), 1/19 * \ell) \qquad \ell = 1, \ldots, 19,$$

and for each ℓ we set $\alpha_{\ell;j,k}^0 = \hat{u}(0, x_\ell)$ for all j, k.

We divided the total population into six hundred eight subpopulations with

$$r_j = 0.3 + 1/3 * 4.7 * (j - 1) \qquad j = 1, 2, 3, 4$$

$$\gamma_1 = 16/38, \ \gamma_2 = 22/38, \ \gamma_3 = 24/38, \tag{20}$$

$$\gamma_k = 16/38 + 1/7 * 22/38 * (k - 1) \qquad k = 4, \ldots, 8.$$

The variables $\alpha_{\ell;j,k}^0$ were then inserted into (19) and the quadratic programming problem was solved for the probability measures p_{jk}^0. The corresponding aggregate population density $u^0(t, x)$ was formed by substituting $\alpha_{\ell;j,k}^0$ and p_{jk}^0 into (12).

These computed solutions $u^0(t, x)$ did not fit the field data as well as any of the previous computed solutions (see Figure 7), returning a residual of $J_3^0 = 0.0020$, and a residual over Days 202, 209 and 216 of $\tilde{J}_3^0 = 0.0018$.

Continuing the iterative process, we used p_{jk}^0 to optimize (14) for $\alpha_{\ell;j,k}^1$. This led to a new set of computed solutions $u^1(t, x)$, which fit the field data better than the previous solution set (see Figure 7), and returned the improved residuals $J_3^1 = 5.6266 \times 10^{-4}$ and $\tilde{J}_3^1 = 2.1284 \times 10^{-4}$. We also found that $|u^0(0, x) - u^1(0, x)|_2 = 0.0600$.

Using $\alpha_{\ell;j,k}^1$ we then optimized (19) for p_{jk}^1. The resulting $u^2(t, x)$ did not appear to improve the fit to the field data, as illustrated by the residuals $J_3^2 = 5.6266 \times 10^{-4}$ and $\tilde{J}_3^2 = 2.1284 \times 10^{-4}$. In addition, we noted that $|\mathbf{p}^0 - \mathbf{p}^1|_2 = 1.9451 \times 10^{-6}$, which also suggested that the latest iteration did
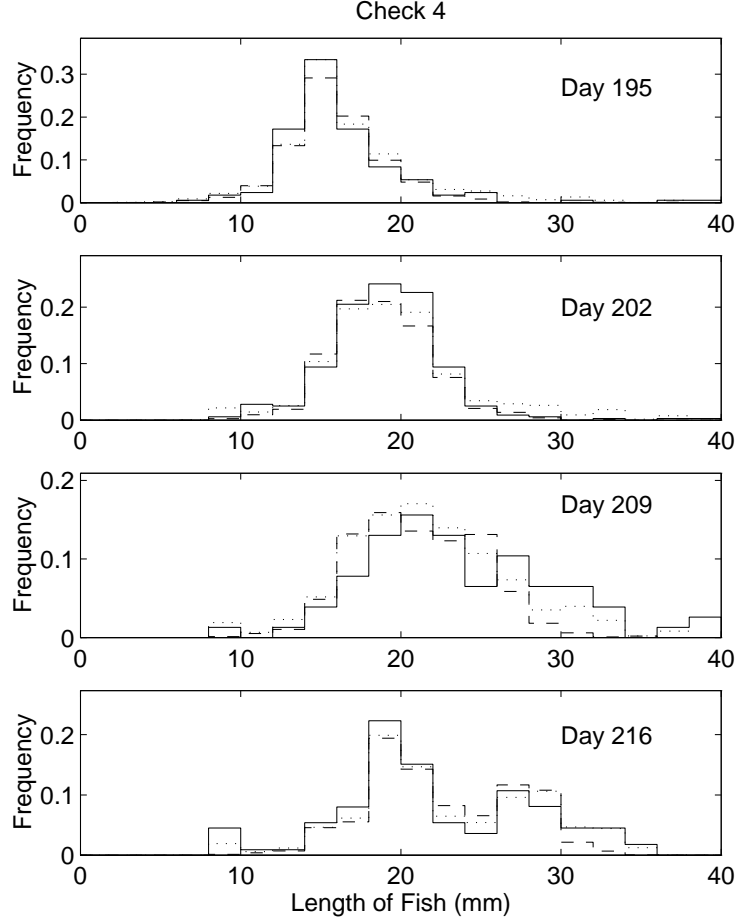
Figure 7: Computed results $u^0(t, x)$ (- -) and $u^1(t, x)$ ($\cdot \cdot$) vs. field data (—) for Example 3.

not bring much improvement. The computed solutions $u^2(t, x)$ are not shown here since the difference between $u^1(t, x)$ and $u^2(t, x)$ are not perceptible to the eye at this resolution.

Finally, we used p^1_{jk} to optimize (14) for $\alpha^2_{\ell;j,k}$. The resulting fit to the field data did improve slightly over the previous fit, returning the residuals $J^3_3 = 4.6647 \times 10^{-4}$ and $\tilde{J}^3_3 = 1.4301 \times 10^{-4}$. We also found that $|u^2(0, x) - u^3(0, x)|_2 = 0.0043$. Even though there was an improvement in the residuals, there was no perceptible improvement visually in the graph of $u^3(t, x)$ compared to $u^2(t, x)$.

The probability densities $p^0(r, \gamma)$ and distribution $P^0(r, \gamma)$ are shown in Figures 8 and 9 respectively. The densities and distributions for $\mathbf{p}^1$ are not shown here because the difference between $\mathbf{p}^0$ and $\mathbf{p}^1$ is not perceptible to the eye at

16

this resolution. Note that for fixed r, the probability distribution $P(\gamma)$ suggests a bimodal population density, as obtained in [BF]. The probability densities $p(r, \gamma)$ do not contribute to this conclusion, and in fact limiting densities themselves may not exist for measures in the set $\{P(g) : g \in G\}$ for general infinite dimensional families G. (see [BF]; the *distributions, not the densities*, can be approximated as we have done in these computations.)
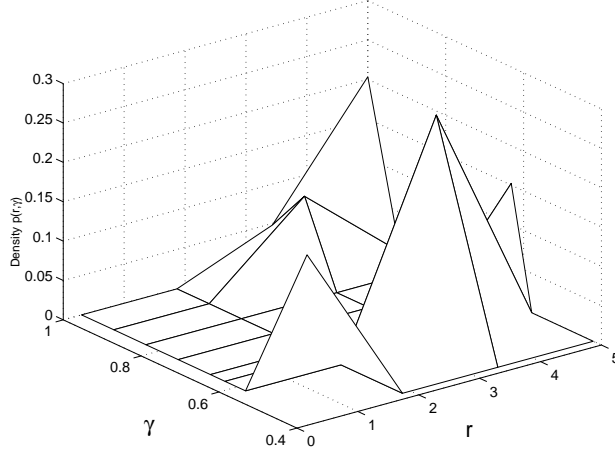

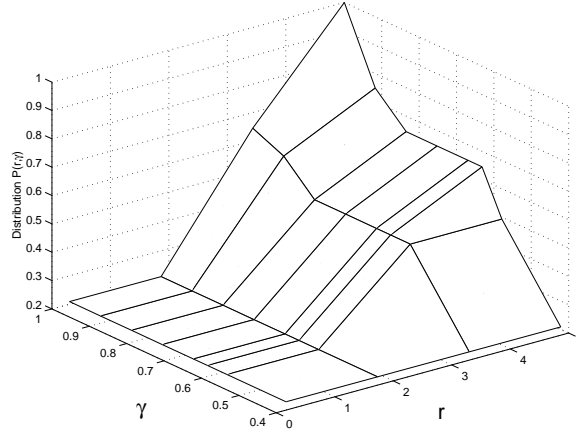Figure 8: Probability densities p_{jk}^0 for Example 3.

Figure 9: Probability distribution P_{jk}^0 for Example 3.

3   Acknowledgments

References

[B1]   H.T. Banks, *Computational techniques for inverse problems in size-structured stochastic population models*, Proceedings IFIP Conference on Optimal Control of Systems Governed by PDE, (Santiago de Compostela, July, 1987) Springer LN in Control and Info. Science, **114** (1988), 3-10.

[B2]   H.T. Banks, *Some remarks on estimation techniques for size-structured population models*, CRSC-TR92-11, NCSU; in Frontiers of Theoretical Biology, (S. Levin, ed.) Springer-Verlag, Berlin, Lecture Notes in Biomathematics, **100** (1994), 609-623.

[BBKW] H.T. Banks, L.W. Botsford, F. Kappel and C. Wang, *Modeling and estimation in size-structured population models*, LCDS-CCS Report 87-13, Brown University; Proceedings 2nd Course on Mathematical Ecology, (Trieste, December 8-12, 1986) World Press, Singapore, 1988, 521-541.

[BF]   H.T. Banks and B.G. Fitzpatrick, *Estimation of growth rate distributions in size-structured population models*, Quart. Appl. Math **49** (1991), 215-235.

[BFZ]   H.T. Banks, B.G. Fitzpatrick, and Y. Zhang, *Estimation of distributed individual rates from aggregate population data*, CRSC-TR94-13, NCSU; in Differential Equations and Applications to Biology and to Industry, ed. by M. Martelli, et al, World Science Press, 1996, 13-22.

[BTW]   H.T. Banks, H.T. Tran, and D.E. Woodward, *Estimation of variable coefficients in the Fokker-Planck equations using moving node finite elements*, CAMS Tech. Rep. 90-9, August 1990, University of Southern California; SIAM J. Num. Anal. **30** (1993), 1574-1602.

[BVWLKRC] L.W. Botsford, B. Vandracek, T. Wainwright, A. Linden, R. Kope, D. Reed and J.J. Cech, *Population development of the mosquitofish, Gambusia Affinis, in rice fields*, Environmental Biology of Fishes, **20** (1987), 143-154.

[F1]   B.G. Fitzpatrick, *Modeling and estimation problems for structured heterogeneous populations*, Journal of Mathematical Analysis and Applications, **172** (1993), 73-91.

[F2]   B.G. Fitzpatrick, *Rate distribution modeling for structured heterogeneous populations*, CRSC-TR93-20, NCSU; in Control and Estimated Parameter Systems: Nonlinear Phenomena, (in Desch, et al, eds.), Intl. Series in Num. Math., Birkhäuser, **118** (1994), 131-141.

[FL1]   W.H. Fleming, *Distributed Parameter Stochastic Systems in Population Biology*, in Control Theory, Numerical Methods and Computer Systems Modelling, (M. Bechmann and H.P. Künzi, eds.) Springer-Verlag, Berline, Lecture Notes in Economics and Mathematical Systems **107** (1974), 180-191.

[FL2]   W.H. Fleming, *A selection-migration model in population genetics*, J. Math Biology **2** (1975), 219-233.

[FS]   W.H. Fleming, and C.H. Su, *Some one-dimensional migration models in population genetics theory*, Theor. Population Biology **5** (1974), 431-449.

[H]   G.W. Harrison, *Numerical solution of the Fokker-Planck equation using moving finite elements*, Num. Meth. for PDE, **4** (1988), 219-232.

[IKP]   K. Ito, F. Kappel and G. Peichl, *A fully discretized approximation scheme for size-structured population models*, Siam J. Numerical Analysis **28** (1991), 923-954.

[O]   A. Okubo, *Diffusion and Ecological Problems: Mathematical Models*, Springer-Verlag, New York, 1980.

[W]   G.H. Weiss, *Equations for the age structure of growing population*, Bull. Math. Biophys., **30** (1968), 427-435.